

Fault Diagnosis with Orthogonal Compactors in Scan-Based Designs

Brady Benware · Grzegorz Mrugalski · Artur Pogiel ·
Janusz Rajski · Jędrzej Solecki · Jerzy Tyszer

Received: 15 October 2010 / Accepted: 10 August 2011 / Published online: 26 August 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract This paper presents a novel scheme to address the challenge of identifying failing scan cells from production test responses in the presence of scan compaction. The scheme is based on a very simple test response compactor employing orthogonal—spatial and time—signatures. The advantage of this scheme as compared to previous work in this field is the simple and incremental nature of the compaction hardware required. The ability of the scheme to accurately identify failing scan cells from compacted responses has been measured on production fail data from five industrial designs and is reported herein.

Keywords Bipartite graphs · Fault diagnosis · Scan-based designs · Test response compaction

Responsible Editor: C. Metra

B. Benware · G. Mrugalski · A. Pogiel · J. Rajski
Mentor Graphics Corporation,
Wilsonville, OR 97070, USA

B. Benware
e-mail: brady_benware@mentor.com

G. Mrugalski
e-mail: grzegorz_mrugalski@mentor.com

A. Pogiel
e-mail: artur_pogiel@mentor.com

J. Rajski
e-mail: janusz_rajski@mentor.com

J. Solecki · J. Tyszer (✉)
Poznań University of Technology,
60-965 Poznań, Poland
e-mail: tyszer@et.put.poznan.pl

J. Solecki
e-mail: jedrzej.solecki@gmail.com

1 Introduction

Failing scan cell data collected from scan-based designs has successfully been used for many years to help identifying systematic yield issues. First of all, failing scan cell knowledge can be used directly, for example to understand where errors are captured within the design and to identify devices, which are defective in a similar region of the chip. Second, localization of failing gates and nets can be obtained through fault diagnosis. From this more precise localization and additional characteristics, systematic defects can be identified. With scan compaction becoming a new production test standard, failing scan cells are no longer directly observable. Well-established techniques are available to perform fault diagnosis from the compacted responses [18]. However, for some failure mechanisms (for example a power droop issue, which produces interacting faults), analyzing failing scan cell data may be the preferred method over fault diagnosis. In these situations, being able to derive the failing scan cells from compacted test responses can be an important capability.

In principle, diagnostic algorithms are classified according to response compaction schemes they rest on. Three basic classes of compaction devices, i.e., infinite input response (time) compactors, finite input response (convolutional) compactors, and combinational (space) compactors drive relevant groups of diagnostic methods. Many of these techniques were developed for built-in self-test (BIST) applications, and they are well documented.

The works of [21] and [28] were first to use signature analysis for diagnostic purposes. They were quickly followed by a similar technique [3] based on multiple-input signature registers (MISR). Other schemes capable of identifying up to a pre-specified number of errors in a similar environment were presented in [4], [16], and [30]. A

finite state machine, inserted between scan chains and a MISR, can be used to diagnose one scan chain at a time, as shown in [31]. It is also possible to arrive with characteristic polynomials for signature registers such that any specified number of errors can be identified [6].

A common approach is to re-run test in order to diagnose more accurately a larger number of scan errors. For instance, a hybrid compactor of [9] combined with a pruning technique has such an ability. A scheme based on Reed-Solomon codes and a programmable MISR [33] collects signatures during BIST sessions, repeated for each feedback polynomial. A set of non-linear equations is subsequently solved to identify the set of failing scan cells. A partitioning-based method that uses an LFSR to gate inputs of a MISR-based compactor was proposed in [26]. For each BIST session, a different pseudorandom selection of scan cells is used. Test time and diagnostic resolution of this method was improved in [1] by introducing deterministic partitioning of scan cells. The problem of identifying failing scan cells can also be mapped into a binary search [10] by deploying extra hardware to produce diversity of deterministic partitions of observed scan cells. In order to improve diagnostic resolution, a two-step scan chain partitioning [20] maintains the principle of random selection of scan chains, but generates additional partitions comprising groups of consecutive scan cells.

Results of space compaction can be essential to the scan-based diagnosis. Combinational compactors, however, must observe each scan chain on two or more outputs, and this is usually done in a per-scan-cycle diagnosis mode. The X-Compact [22] can uniquely identify a single failing scan chain provided that only one scan chain produces an error at any scan-out cycle. Other failures need an exhaustive checking guided by statistical data [29]. The direct diagnosis of [13] constructs a per-cycle combinational representation of a circuit under test and employs simple parity information to arrive with fault-model independent diagnostic results. The i-Compact scheme [25] employs results of distance-based coding theory to diagnose various error combinations. An XOR network-based compactor complements a BIST environment in the SDBIST scheme proposed in [32]. Any single-chain output failure is uniquely diagnosable in this approach as every scan chain is connected to a different subset of outputs.

The *convolutional* compaction has the ability to uniquely encode a test response of several scan-out cycles into a signature that can be observed in more than one scan shift cycle. As a result, it has a potential for accurate fault diagnosis in scan-based designs. For example, the method of [23] uses a branch-and-bound algorithm to narrow the set of scan cells down to the sites that are most likely to capture faulty signals. This search is guided by a number of heuristics and self-learned information used to accelerate

the diagnosis for the subsequent test patterns. Furthermore, this approach belongs to a class of solutions that attempt to identify failing scan cells even in the presence of unknown (X) states. The i-Compact [25] allows identification of errors in the presence of X states by following the concept of erasures in the coding theory. The scheme of [19] uses an LFSR to randomly select scan chains whose contents are then XOR-ed to produce a parity bit observed every scan-out cycle. The X-Compact can also be used to identify failing scan chains in the presence of X states [29]. The diagnostic technique is quite similar to that of [22], but in order to reduce the impact of X states, one has to increase the number of outputs. Finally, various specialized techniques have been proposed to diagnose scan chain failures [7], [11], [14], [17], [24].

This paper presents a novel *indirect* fault diagnosis technique that takes advantage of a very simple compaction scheme and can be used in high-volume production diagnosis. We demonstrate that it is possible to accurately and quickly identify failing scan cells by using companion signatures obtained in parallel through a conventional space compaction and a simplified finite input response filter. The essence of the method is to find a maximum bipartite matching in a graph representing compacted test results. This process is guided by heuristics designating the most likely sites of failing scan cells. Experimental results include numbers obtained for industrial designs and actual fail log information collected during production scan testing.

2 Compaction Scheme

Figure 1 illustrates an example of an orthogonal compaction scheme working with a number of scan chains and having two outputs. As can be seen, the compaction circuitry comprises both combinational and sequential parts connected to the outputs of scan chains. The spatial part of the compactor reuses a single-output XOR network (a parity tree) commonly deployed in a variety of scan-based test architectures. The same scan outputs are further connected to a shift register through individual XOR gates interspersed between its memory elements. There is no feedback loop in the register. As a result, all scan cell errors can produce an m -bit spatial signature (one bit per scan shift cycle), where m is the length of the longest scan chain. Moreover, the same errors yield an n -bit time signature, where $n = m + s - 1$, and s is the number of scan chains.

In this work, we assume, moreover, that a selective scan chain masking mechanism shapes an X-state profile of a CUT as shown in Fig. 1. It can be done in such a way that all unknown states are suppressed before they enter the compactor. This is achieved by using a reasonable amount of control data and without compromising test coverage and

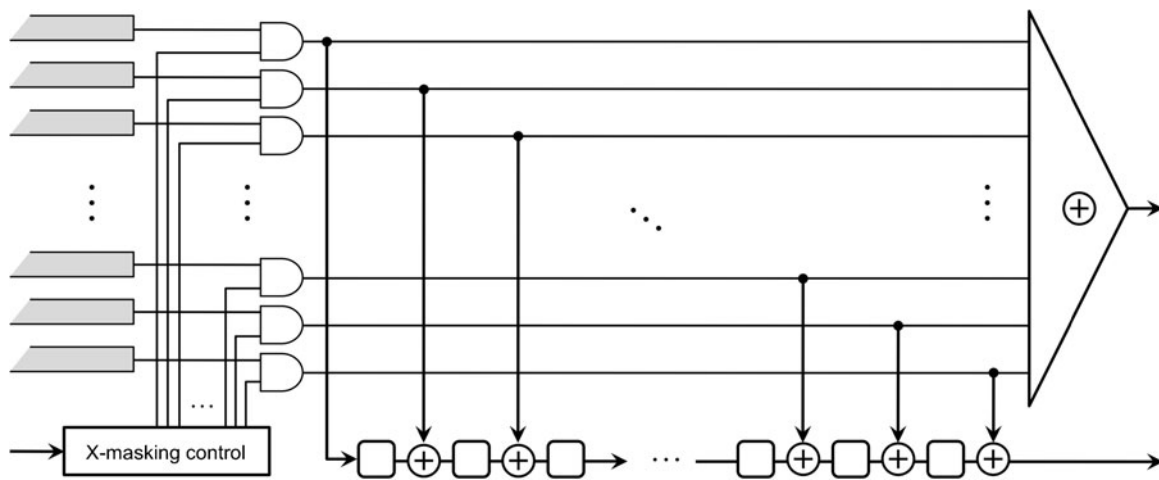


Fig. 1 Orthogonal test response compaction

test data compression, as shown in our earlier works. In particular, the X-masking can be carried out in either per pattern [27] or per cycle mode [5], with the latter solution allowing all test patterns to share a single set of control data obtained from aggregated X-profiles.

Since the compactor is a linear circuit, we will analyze its behavior using error test responses it receives and error signatures it produces. An error test response is a bit-wise XOR of fault-free and faulty test responses, respectively. Similarly, an error signature is a sum of fault-free and faulty signatures. In the rest of the paper, we will consider only error test responses and error signatures.

Producing two companion signatures is essentially carried out as shown in Fig. 2. The upper binary sequence represents a spatial signature that is generated by the single 6-input XOR tree. The bottom sequence corresponds to the content of the time signature. The figure demonstrates a few possible scenarios according to which errors either can mask each other in space (if they occur in scan cells

belonging to the same scan shift cycle or, alternatively, time slice) or in a time domain provided they are caught by scan cells located on conceptual diagonals. Clearly, errors canceled in one signature have still a good chance to survive in the other one. This increases likelihood of both fault detection and its accurate identification.

It is worth noting that a given time signature is typically longer than the longest scan chain, and the size of its “tail” depends on the number of scan chains. In all diagnostic procedures discussed in the following, we assume that test results collected during application of consecutive test patterns and represented by time signatures may overlap. In particular, bits forming a tail of the previous signature are XOR-ed with data captured in a few first slices of scan chains and corresponding to the next test response. Possible mutual interactions between erroneous bits of both responses have to be taken into account when running diagnostic methods, as demonstrated in Sections 3 and 4.

3 Diagnostic Algorithm

Given two error signatures produced by the orthogonal compaction scheme of the previous section, locations of failing scan cells are determined by using a two-phase process. First, we construct a bipartite graph whose vertices are divided into two disjoint sets **C** and **D** representing time frames (columns) and scan cell diagonals, respectively, for which errors were recorded in the corresponding signatures, as shown in Fig. 3. Every edge in this particular graph connects a vertex *c* in **C** with a vertex *d* in **D** provided column *c* intersects diagonal *d* within scan chains, which are the subject of diagnosis.

Consider six scan chains, each 10-bit long, shown in Fig. 3. Let a fault propagate to scan cells indicated by black dots, thus forming an error of multiplicity four. Additional

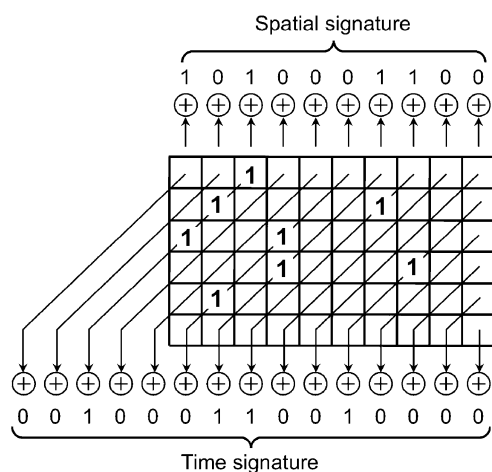


Fig. 2 Computing spatial and time signatures

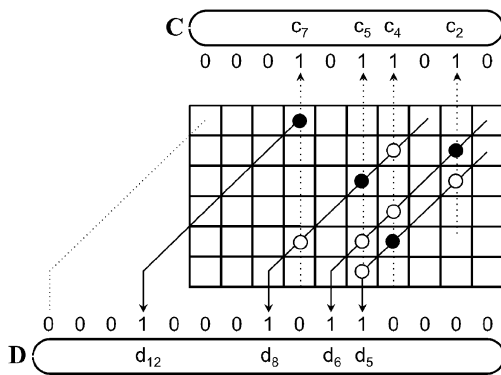


Fig. 3 Example of error pattern

intersections of those diagonals and columns that yield ones in the error signatures are depicted by white dots. As a result, diagonals d_5, d_6, d_8, d_{12} as well as columns c_2, c_4, c_5 , and c_7 form a bipartite graph presented in Fig. 4.

Identifying of failing scan cells is now equivalent to finding a matching M in the bipartite graph (similar to that of Fig. 4), i.e., determining an edge set such that no two edges of M share their endpoints. Note that every matching edge uniquely indicates a failing scan cell. Clearly, the ultimate objective of this phase is to arrive with a *perfect bipartite matching*, that is, a solution that matches all vertices in the graph (in other words every vertex is incident to exactly one edge of the matching). Stated differently, any perfect matching we can obtain (if possible) is a maximum one having the largest number of edges. As can be seen, the set of matching edges denoted by bold lines in Fig. 4 corresponds to the original error of Fig. 3.

In general, different errors can cause the same error signatures. Indeed, when injecting errors into a signature, scan cells interact with each other on the compactor register and inputs of the XOR tree. Consequently, certain error signals can be masked, thus leading to ambiguity in selection of the actual scan sites catching the errors. Even if there is no error masking, there are still different errors, which are equivalent causes of the recorded signatures. The latter phenomenon is illustrated in Fig. 5 where another perfect matching in the bipartite graph indicates a different error pattern yielding the same signatures as those of Fig. 3.

In addition to the cases in which the number of ones in error signatures is equal to the number of scan cells affected

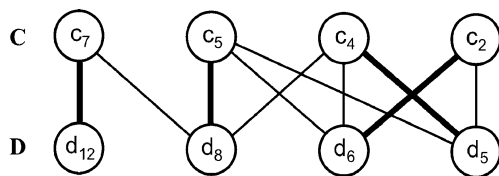


Fig. 4 Bipartite graph for error of Fig. 3

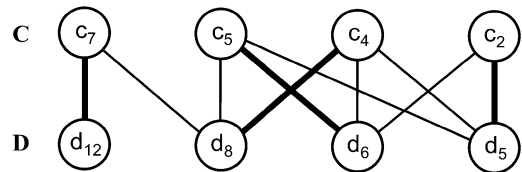


Fig. 5 Perfect matching for error of Fig. 3

by a fault, the diagnostic scheme has to handle scenarios with error masking. In principle, there are three of them: (1) an even number of ones occurring in a given time frame (scan shift cycle) leads to a spatial aliasing, (2) an even number of ones is observed along one or more diagonals causing an error cancellation on the corresponding segments of the shift register, (3) both phenomena occur at the same time.

For the sake of illustration, consider an error shown in Fig. 6. As can be verified, a vertical alignment of two scan cells catching errors causes spatial aliasing. Hence, working with the resultant error signatures yields a bipartite graph as depicted in the same figure. Clearly, there is no perfect matching in this graph. Instead, one can determine a *maximum matching*, i.e., a matching that contains the largest possible number of edges. However, the time signature content reveals that a solution consists of more failing cells than determined by any maximum matching. These additional failing sites cannot be uniquely determined without additional data, since they may occur in multiple locations along the suspect diagonals. This problem is resolved by using test responses collected for the remaining test patterns, as shown in Section 4. Furthermore, the above example justifies an initial selection of a solution multiplicity. Given the number of ones in a spatial signature and in a time signature, respectively, we

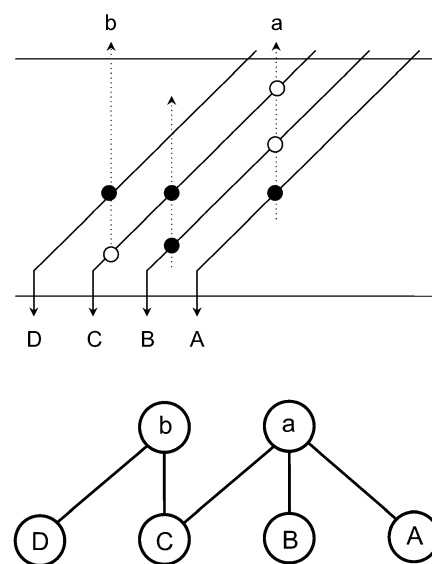


Fig. 6 Spatial error masking

limit the number of failing scan cells to the higher value of the two.

A scenario similar to that of Fig. 6 is also observed for error masking occurring either exclusively in the shift register or in both time and space dimensions. Again, only a maximum matching can be found, leaving further steps for a diagnostic process that takes advantage of test results collected for the remaining test patterns.

When time signatures are created for conceptual diagonals of scan cells, some of these diagonals traverse successive error patterns, causing effectively signature overlapping, as already mentioned in Section 2. Hence, the current error pattern is always processed in conjunction with its two neighbors in such a way that both overlapping parts are included in the diagnostic analysis by treating them as integral parts of one error pattern. This phenomenon is illustrated in Fig. 7, where the actual time signature is obtained by overlapping test responses from three consecutive responses, and is accompanied by the corresponding bipartite graph.

Once the bipartite graph representing both recorded signatures is formed, we use the Hopcroft-Karp algorithm [12] to obtain the first solution, i.e., a maximum bipartite matching representing a set of scan cells that capture errors and then produce both signatures. The algorithm repeatedly increases the size of an initial partial matching by finding a maximal set of shortest augmenting paths. An augmenting path starts at a free vertex, i.e., a non-end-point of an edge in some partial matching M , ends at a free vertex, and alternates between unmatched and matched edges within the path. If M is a matching of size n , and P is an augmenting path relative to M , then the set $M \oplus P$ forms a matching with size $n+1$. Thus, by finding augmenting

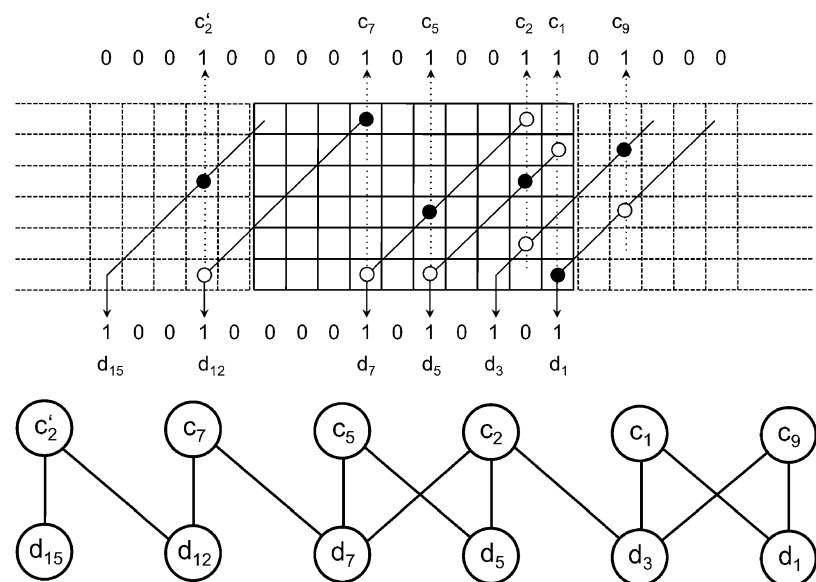
paths, the algorithm increases the size of the matching. As a result, it runs in $O(e\sqrt{v})$ time in the worst case, where e is the number of edges, and v is the number of vertices.

4 Selection Heuristics

As shown in the previous section, the diagnostic procedure begins by using the Hopcroft-Karp algorithm. However, given an inherent ambiguity of mapping of actual errors into signatures, one has to carry on in order to respond to a possibility of having many equivalent solutions for a single test pattern. In order to determine such error patterns, we need to find all maximum matchings in a given bipartite graph.

In our approach, we adapt the algorithm of [8]. Given a bipartite graph G and an initial matching M , we orient edges of M from C to D (see Section 3) and other edges in the opposite direction. Then, a depth-first search finds a cycle δ which is used to produce a new matching M' by exchanging edges along the cycle, i.e., $M' = M \oplus \delta$. For example, the symmetric difference of cycle $\delta = d_5 \rightarrow c_2 \rightarrow d_6 \rightarrow c_5 \rightarrow d_8 \rightarrow c_4 \rightarrow d_5$ and the perfect matching of Fig. 4 yields the perfect matching of Fig. 5. Subsequently, we select an edge $e \in M - M'$, and create two sub-graphs G_1 and G_2 of G . G_1 is obtained by deleting all edges adjacent to e , whereas G_2 is obtained by just deleting e from G . These two sub-graphs, as a result of recursive calls, become now the subject of the same procedure invoked with the initial matchings M and M' as input parameters, respectively. Consequently, a binary tree represents generation of successive maximum matchings with each node corresponding to one solution. During this process, we maintain individual counters associated with each scan cell. Given a scan cell x

Fig. 7 Test response overlapping in time signature



and a test pattern t , the value of counter $v_t(x)$ is equal to the number of occurrences of cell x in all error patterns identified for a test vector t . Its value may be regarded as proportional to the probability that scan cell x belongs to one or more actual error patterns. Finally, given the number E_t of error patterns found for test t , we assign a weight $w(x)$ to cell x , where $w(x)$ is defined as a maximum over all values $100 \cdot v_t(x)/E_t$.

The above weights guide the selection process to arrive with the most likely locations of failing scan cells as follows. For each error pattern, we determine its score by summing the corresponding scan cell weights. Moreover, the score is normalized with respect to the number of scan cells and reduced by the number of scan chains hosting involved scan cells (it gives higher priority to solutions occupying fewer scan chains). An error pattern with the largest score is then the final solution.

Consider two test vectors and scan cells A, B, C, D, E. Let two sets of error patterns $\{AB, AC\}$ and $\{ABC, ABD, ABE\}$ be associated with these vectors. Then we have: $v_1(A)=2$, $v_1(B)=1$, $v_1(C)=1$, $v_2(A)=3$, $v_2(B)=3$, $v_2(C)=1$, $v_2(D)=1$, $v_2(E)=1$, as shown in Fig. 8. Hence, $w(A)=\max\{100 \cdot 2/2, 100 \cdot 3/3\}=100$, $w(B)=\max\{100 \cdot 1/2, 100 \cdot 3/3\}=100$, $w(C)=50$, $w(D) \approx 33$, $w(E) \approx 33$. Assuming that only cells B and E share the same scan chain, the ranking of the error patterns is as follows: AB $\{(100+100)/2-2=98\}$, AC $\{(100+50)/2-2=73\}$, ABC $\{(100+100+50)/3-3 \approx 80\}$, ABD $\{(100+100+33)/3-3 \approx 75\}$, ABE $\{(100+100+33)/3-2 \approx 76\}$. Hence, we choose errors AB and ABC as the most likely solutions.

If there are two or more error patterns with the same rank, two additional criteria are taken into account. Given a scan cell x , it receives a weight equal to the number of test patterns for which x has been designated by our diagnostic method as a cell receiving an error. Also, a scan chain

hosting x is assigned a similar metric based on its presence in other solutions. The score of a given error pattern is then determined by summing the corresponding scan cell weights, and, if it does not break a tie, by summing weights characterizing the corresponding scan chains.

With the increasing complexity of scan and errors, the tree of sub-problems grows quickly, thus leading to numerous solutions. Therefore, we first obtain solutions for simpler errors and then reiterate over more complicated cases to prune candidates with cells unlikely to catch errors. Indeed, many experimental evidences show that for different stimuli a given fault propagates to the same scan cells, their neighbors, or to the same scan chains. One can take advantage of this observation by preferring cells located in scan chains hosting cells already considered as likely sites of errors, and by designating explicitly certain scan chains as the most likely locations of failing scan cells.

Producing all matchings for larger errors is infeasible because of time. Fortunately, these error patterns can be analyzed using a weighted matching algorithm. For a given error pattern, we create a bipartite graph, as shown earlier. Next, the algorithm assigns weights to its edges. The weights are dependent on diagnostic results obtained for smaller errors. For a given graph, we determine then a maximum weight matching by using a method of [15] to solve the assignment problem in a polynomial time. Such a maximum matching becomes the final diagnosis outcome.

The following steps summarize the diagnostic procedure (for one test session) proposed in this section:

```

for each recorded signature create the bipartite graph
for each graph
    if the error multiplicity is large
        find a maximum weight matching
        output the result and exit
    else
        find all maximum matchings as candidate solutions
        rate each scan cell within a given candidate solution
for each scan cell assign the maximum rating to it
for each graph
    rate candidate solutions
    output a candidate solution with the maximum rating
  
```

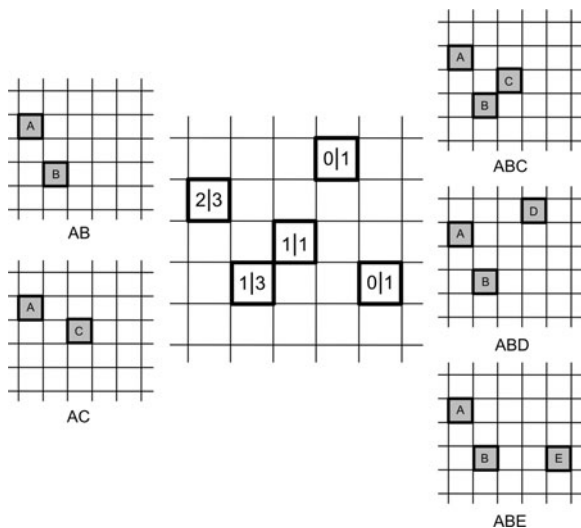


Fig. 8 Scan cell weights

5 Experimental Results

The objective of this experimental analysis was to study feasibility of the proposed scheme. The primary target in all experiments is the ability of the scheme to recreate, based on test responses collected in the production test environ-

ment, original sites of failing scan cells for successive test patterns. The experiments were conducted on several industrial designs on material which represented manufacturing conditions typical of a mature manufacturing environment (i.e., mature yields with occasional excursions). Their characteristics, including the number of gates and scan architecture, are given in Table 1. Also, for each test case, the table provides compaction rates, the number of examined faulty chips and unique erroneous test responses, as well as the number of cases in which a two-dimensional undetectable aliasing (both in space and in time) has occurred. As can be seen, the compaction ratio is obtained in each test case by halving the number of scan chains as we collect signatures on just two outputs. Deploying additional spatial and/or time compactors (to reduce the probability of aliasing in both dimensions) decreases this ratio. The gain would be essentially that of the diagnostic resolution. As demonstrated in the remaining of this section, however, the quality of current results can hardly be used to justify such a solution.

A *diagnostic efficiency* was employed as a figure of merit to assess performance of the scheme. Given a circuit and all unique signatures that capture errors, the diagnostic efficiency is a fraction of error patterns whose affected scan cells are all correctly identified, i.e., there are no missing scan cells or extra scan cells in a solution compared to the actual error. Recall that a test set can detect certain faults several times, and subsets of affected scan cells may differ each time. As discussed in Section 4, we use such information to arrive with a more accurate identification of failing scan cells, although different error patterns are counted separately in tables to follow.

The results with respect to the diagnostic efficiency (DE) for all the examined circuits are shown in Table 2. Each column of the table corresponds to a given multiplicity E of error patterns, i.e., to the number of failing scan cells that constitute the recorded error patterns. Then, for each design, the row $\#S$ reports the number of observed signatures corresponding to error patterns of a given size, while the row $\#S$ [%] provides the same data as a percentage. Recall that a large body of earlier experimental evidence shows that for many stimuli faults propagate to a very few scan cells [23]. This is also confirmed in Table 2,

where presented data were obtained by using fail logs collected during production test. As can be seen, errors of multiplicity up to 10 produce a significant majority of recorded erroneous signatures. The next two rows report the corresponding diagnostic efficiency. The first number (DE) gives the percentage of cases diagnosed correctly, whereas the second quantity reports a cumulative diagnostic efficiency (CDE), i.e., the fraction of successfully identified error patterns of multiplicities up to a value represented by a respective column of the table. As can be seen, for error patterns of size up to 10, the cumulative diagnostic efficiency remains above 96%.

The same table provides data regarding processing time (in milliseconds) needed to arrive with the reported results (rows labeled CPU time). All experiments were carried out on a computer with 2.4 GHz dual core processor and 4.0 GB of RAM. As can be seen, the CPU time for identifying most of the errors is less than one-third of a second. Interestingly, for some designs (D1, D5), the CPU time needed to identify errors of larger multiplicities is actually decreasing. This is due to peculiar error patterns affecting 8 and more scan cells, which are producing signatures with relatively small number of errors (ones).

It is worth noting that typically even solutions not counted in the tables as perfect matches comprise scan cells that belong to the actual error patterns. This trend has been confirmed for all test cases examined in the experiments. It is illustrated in the row (2|2) where we report the percentage of solutions in which, in addition to correctly identified scan cells, the number of missing or extra scan cells does not exceed two. The last row is the sum of the rows DE and 2|2—it clearly shows that the significant majority of errors are covered by cases reported in the tables.

It might be also of interest to see how the selection heuristics, introduced in Section 4, can impact the quality of diagnostic reasoning. Some numbers addressing this concern are presented in Table 3. They have been obtained and averaged out over all benchmark designs used in this paper. Given heuristics based on cell weights (H_1), faulty scan cell counts (H_2), and faulty scan chain counts (H_3), we report—for each error multiplicity from 2 to 8—two figures of merit: the percentage of cases where certain error patterns were correctly rejected (CR) as false

Table 1 Circuit characteristics

	Gates	Scan chains (no×size)	Comp. [x]	Faulty chips	Err. test responses	Alias
D1	22.9 M	799×801	400	1256	21936	93
D2	507 K	160×259	80	988	16532	808
D3	1.06 M	288×290	144	1040	12198	231
D4	1.47 M	200×525	100	334	42439	432
D5	1.09 M	160×470	80	977	3245	0

Table 2 Diagnostic efficiency [%] (CPU time is measured in milliseconds)

<i>E</i>	1	2	3	4	5	6	7	8	9	10
Design D1										
# <i>S</i>	21741	13485	6557	3808	2325	1534	1043	815	648	468
# <i>S</i> [%]	39.56	24.54	11.93	6.93	4.23	2.79	1.90	1.48	1.18	0.85
DE	99.69	98.91	97.21	94.38	92.60	85.72	82.93	79.75	71.45	62.61
CDE	99.69	99.39	99.05	98.66	98.36	97.97	97.66	97.38	97.05	96.75
CPU time	12.15	10.78	18.50	30.13	38.63	86.86	100.87	212.27	240.38	195.00
2 2	0.14	0.55	0.72	1.21	1.12	1.89	1.82	1.10	1.08	1.71
DE+2 2	99.83	99.46	97.93	95.59	93.72	87.61	84.75	80.85	72.53	64.32
Design D2										
# <i>S</i>	2705	1100	1043	1167	1247	1328	1301	1210	953	725
# <i>S</i> [%]	14.94	6.07	5.76	6.44	6.89	7.33	7.18	6.68	5.26	4.00
DE	100.00	99.82	99.71	100.00	99.36	99.40	99.00	98.68	98.53	95.03
CDE	100.00	99.95	99.90	99.92	99.82	99.76	99.66	99.55	99.47	99.22
CPU time	0.76	0.44	0.60	2.27	4.07	6.57	26.52	163.60	153.30	207.40
2 2	0.00	0.00	0.10	0.00	0.64	0.45	0.38	0.25	0.73	0.97
DE+2 2	100.00	99.82	99.81	100.00	100.00	99.85	99.38	98.93	99.26	96.00
Design D3										
# <i>S</i>	8472	4405	2062	1487	1121	980	656	566	335	325
# <i>S</i> [%]	38.62	20.08	9.40	6.78	5.11	4.47	2.99	2.58	1.53	1.48
DE	99.56	98.48	97.96	95.97	94.92	94.90	92.84	90.81	86.57	81.54
CDE	99.56	99.19	99.02	98.75	98.50	98.31	98.12	97.91	97.72	97.47
CPU time	2.73	4.50	10.20	8.39	13.57	31.61	66.25	129.86	177.64	312.37
2 2	0.24	0.66	0.44	1.21	0.98	1.43	2.29	1.41	2.09	2.15
DE+2 2	99.80	99.14	98.40	97.18	95.90	96.33	95.13	92.22	88.66	83.69
Design D4										
# <i>S</i>	8031	8114	7239	7313	5984	4646	3693	2744	2359	1922
# <i>S</i> [%]	12.80	12.93	11.54	11.66	9.54	7.41	5.89	4.37	3.76	3.06
DE	99.90	99.77	99.52	99.21	98.76	98.04	97.16	94.75	93.22	89.91
CDE	99.90	99.83	99.73	99.61	99.47	99.31	99.13	98.88	98.62	98.29
CPU time	7.27	2.82	8.13	4.42	3.79	6.86	9.25	67.11	106.37	185.10
2 2	0.06	0.09	0.23	0.40	0.53	0.73	0.89	2.00	1.70	2.08
DE+2 2	99.96	99.87	99.75	99.62	99.28	98.77	98.05	96.74	94.93	91.99
Design D5										
# <i>S</i>	2563	669	191	157	131	145	56	29	24	17
# <i>S</i> [%]	53.97	14.09	4.02	3.31	2.76	3.05	1.18	0.61	0.51	0.36
DE	100.00	100.00	99.48	100.00	100.00	93.79	100.00	100.00	100.00	100.00
CDE	100.00	100.00	99.97	99.97	99.97	99.74	99.74	99.75	99.75	99.75
CPU time	0.04	0.07	0.74	0.10	0.53	28.22	18.62	57.44	18.11	6.05
2 2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DE+2 2	100.00	100.00	99.48	100.00	100.00	93.79	100.00	100.00	100.00	100.00

solutions, and the percentage of error patterns rejected incorrectly (IR). The latter number is a fraction of the actual errors (failing scan cells) that a given heuristic was unable to recognize and, as shown in the table, it has negligibly small values. The first statistic (CR), on the other hand, illustrates high efficiency of the proposed heuristics in pruning a solution space.

Comparison with prior works The scheme presented in this paper is primarily related to the convolutional test response compaction, and (albeit to the less extent) to various parity-tree-based diagnostic techniques. The latter schemes, however, have to monitor a number of output streams, thus compromising compaction ratios. For example, the X-Compact of [22] can identify a single failing scan chain

Table 3 Efficacy of the selection heuristics [%]

E	H ₁		H ₂		H ₃	
	CR	IR	CR	IR	CR	IR
2	52.90	0.14	52.91	0.14	48.72	0.06
3	70.17	0.12	70.26	0.13	57.69	0.51
4	83.42	0.11	83.74	0.11	74.02	0.76
5	90.30	0.09	90.46	0.08	85.17	0.89
6	93.93	0.07	93.74	0.06	91.28	0.93
7	96.10	0.04	95.19	0.03	94.50	0.95
8	97.88	0.06	96.65	0.04	97.11	0.96

provided it is the only scan chain that produces an error at a given scan-out cycle. Other failures need an exhaustive checking. Similar rules govern SDBIST-based fault diagnosis [32] and the use of coding-theory results, as demonstrated by the i-Compact technique [25].

Table 4 compares results presented in this paper with diagnostic efficiency of a branch-and-bound algorithm [23] working with test responses produced by 64-bit convolutional compactors (except design D1 where we have used a 96-bit compactor). Given a design and an error multiplicity listed in the first column, each entry to the table provides a difference between diagnostic efficiency reported in Table 2 and that of the convolutional diagnosis. Thus, all positive items in Table 4 indicate superiority of the new scheme. As can be observed, the ability of both schemes to recognize various error combinations varies depending on the error cardinality. Typically, the new scheme offers visibly better performance, which is particularly appealing for designs D4 and D5. These results make the new approach a very attractive option given its much simpler hardware. In a few test cases, we observe a marginal superiority of the convolutional diagnosis. It can be attributed to its more sophisticated injector network that allows easier or faster identification of some of the error patterns.

Table 4 Comparison with the convolutional compaction [%]

E	D1	D2	D3	D4	D5
1	0	0	0	0	0
2	0.05	−0.12	−0.24	2.87	0
3	−0.03	−0.11	0.07	6.29	−0.56
4	0.12	0.38	2.5	7.09	0
5	0.83	−0.34	4.17	9.44	0
6	0.15	−0.15	5.59	9.78	1.75
7	3.32	−0.78	5.5	12.7	2.17
8	1.4	−0.84	1.95	13.13	14.81
9	2.55	−2.12	−3.9	14.78	42.86
10	4.26	0.14	0.23	14.9	87.5

Hardware overhead The silicon area of the proposed compaction circuitry (as shown in Fig. 1) amounts to a certain number of 2-input XOR gates and flip-flops. Table 5 provides the actual area costs computed with a commercial synthesis tool for the designs of Table 1 (except D5 whose EDT logic is similar to that of D2). All components of our solution were synthesized as an integral part of EDT compression logic by using a 90 nm CMOS standard cell library under 10 ns timing constraint. The table reports the resultant EDT logic [27] silicon overhead as a total equivalent area of two-input NAND gates, and then with respect to combinational (*Comb*) and non-combinational (*Non-Comb*) devices. The area taken by the orthogonal compactor is reported in the second part of the table in rows *Spatial* (the combinational part) and *Time* (the sequential part with all associated XOR gates). Note that the time compactor area is further divided into combinational (*Time-C*) and non-combinational (*Time-NC*) parts, as shown in two subsequent rows of the table. These numbers are then compared with the corresponding EDT area in rows *Spatial* % and *Time*%, respectively, where the area occupied by the compactor is expressed as a percentage of the total EDT area. As can be easily verified, the area overhead of circuitry newly added to EDT for diagnostic purposes, i.e., a shift register with 2-input XOR gates, ranges between 23% and 27% of the entire EDT real estate.

6 Conclusion

In this paper, we tackle the fault diagnosis technique for scan-based designs. It is based on results of orthogonal test response compaction and methods used to find maximum bipartite matchings. The proposed solution matches very well the requirements of embedded deterministic test and support high quality test by providing ability to identify failing scan cells directly from the compacted test responses

Table 5 Area overhead

	D1	D2	D3	D4
EDT	20386	4706	7847	5675
Comb	15950	3494	5994	4262
Non-Comb	4436	1212	1853	1413
Spatial	2298	457	825	572
Spatial%	11.3%	9.7%	10.5%	10.1%
Time	5514	1098	1983	1375
Time-C	1692	337	608	422
Time-NC	3822	761	1375	953
Time%	27.0%	23.3%	25.3%	24.2%

with a minimal impact on existing embedded test logic. Results of experiments conducted on several industrial designs prove feasibility of the proposed approach. In particular, we have shown that even with high compression ratios, exceeding 100x, it is possible to identify exact locations where the errors come from. The scheme is also consistent with the multi-site testing methodology, as it requires only two output pins.

Acknowledgment We presented an earlier version [2] of this paper at the European Test Symposium 2010.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Bayraktaroglu I, Orailoglu A (2005) The construction of optimal deterministic partitioning in scan-based BIST fault diagnosis: mathematical foundations and cost-effective implementations. *IEEE Trans Comput* 54(1):61–75
- Benware B, Mrugalski G, Pogiel A, Rajski J, Solecki J, Tyszer J (2010) “Diagnosis of failing scan cells through orthogonal response compaction.” *Proc. ETS*. 221–226
- Chan JC, Abraham JA (1990) “A study of faulty signatures using a matrix formulation.” *Proc. ITC*. 553–561
- Chan JC, Womack BF (1990) “A study of faulty signature for diagnostics.” *Proc. ISCAS*. 2701–2704
- Czys D, Mrugalski G, Mukherjee N, Rajski J, Tyszer J (2010) On compaction utilizing inter and intra correlation of unknown states. *IEEE Trans Computer-Aided Design* 29(1):117–126
- Damarla T, Stroud CE, Sathaye A (1995) Multiple error detection and identification via signature analysis. *Journal of Electronic Testing: Theory and Applications* 7(3):193–207
- Edirisooriya S, Edirisooriya G (1995) “Diagnosis of scan path failures.” *Proc. VTS*. 250–255
- Fukuda K, Matsui T (1994) Finding all the perfect matchings in bipartite graphs. *Applied Mathematics Letters* 7(1):15–18
- Ghosh-Dastidar J, Das D, Toubia NA (1999) “Fault diagnosis in scan-based BIST using both time and space information.” *Proc. ITC*. 95–102
- Ghosh-Dastidar J, Toubia NA (2000) “A rapid and scalable diagnosis scheme for BIST environments with a large number of scan chains.” *Proc. VTS*. 79–85
- Guo R, Venkataraman S (2001) “A technique for fault diagnosis of defects in scan chains.” *Proc. ITC*. 268–277
- Hopcroft JE, Karp RM (1973) An $n^{5/2}$ algorithm for maximum matchings in bipartite graphs. *SIAM Journal on Computing* 2(4):225–231
- Holst S, Wunderlich H-J (2009) “A diagnosis algorithm for extreme space compaction.” *Proc. DATE*. 1355–1360
- Huang Y, Cheng W-T, Rajski J (2005) “Compressed pattern diagnosis for scan chain failures.” *Proc. ITC*. paper 30.3
- Jonker R, Volgenant A (1987) A shortest augmenting path algorithm for dense and sparse linear assignment problems. *Computing* 38:325–340, Springer-Verlag
- Karpovsky MG, Chaudhry SM (1993) Design of self-diagnostic boards by multiple signature analysis. *IEEE Trans Comput* 42(9):1035–1044
- Kundu S (1993) “On diagnosis of faults in a scan-chain.” *Proc. VTS*. 303–308
- Leininger A, Muhmenthaler P, Cheng W-T, Tamarapalli N, Yang W, Tsai H (2005) “Compression mode diagnosis enables high volume monitoring diagnosis flow.” *Proc. ITC*. 156–165
- Lin H-M, Li JC-M (2005) “Column parity and row selection (CPRS): a BIST diagnosis technique for multiple errors in multiple scan chains.” *Proc. ITC*. paper 42.3
- Liu C, Chakrabarty K (2004) Identification of error-capturing scan cells in scan-BIST with applications to System-on-Chip. *IEEE Trans Computer-Aided Design* 23(10):1447–1459
- McAnney WH, Savir J (1987) “There is information in faulty signatures.” *Proc. ITC*. 630–636
- Mitra S, Kim KS (2004) X-Compact: an efficient response compaction technique. *IEEE Trans Computer-Aided Design* 23(4):421–432
- Mrugalski G, Pogiel A, Rajski J, Tyszer J, Wang C (2007) Fault diagnosis with convolutional compactors. *IEEE Trans Computer-Aided Design* 26(8):1478–1494
- Narayanan S, Das A (1997) “An efficient scheme to diagnose scan chains.” *Proc. ITC*. 704–713
- Patel JH, Lumetta SS, Reddy SM (2003) “Application of Saluja-Karpovsky compactors to test responses with many unknowns.” *Proc. VTS*. 107–112
- Rajski J, Tyszer J (1999) Diagnosis of scan cells in BIST environment. *IEEE Trans Comput* 48(7):724–731
- Rajski J, Tyszer J, Mrugalski G, Mukherjee N, Cheng W-T, Kassab M (2008) X-Press: two-stage X-tolerant compactor with programmable selector. *IEEE Trans Computer-Aided Design* 27(1):147–159
- Savir J, McAnney WH (1988) “Identification of failing tests with cycling registers.” *Proc. ITC*. 322–328
- Stanojevic Z, Guo R, Mitra S, Venkataraman S (2005) “Enabling yield analysis with X-Compact.” *Proc. ITC*. 726–734
- Stroud CE, Damarla T (1995) “Improving the efficiency of error identification via signature analysis.” *Proc. VTS*. 244–249
- Tekumalla RC (2003) “On reducing aliasing effects and improving diagnosis of logic BIST failures.” *Proc. ITC*. 737–744
- Wohl P, Waicukauski J, Patel S, Hay C, Gizdarski E, Mathew B (2005) “Hierarchical compactor design for diagnosis in deterministic logic BIST.” *Proc. VTS*. 359–365
- Wu Y, Adham S (1999) Scan-based BIST fault diagnosis. *IEEE Trans Computer-Aided Design* 18(2):203–211

Brady Benware received Ph.D. degree in electrical engineering from Colorado State University in 2001. He is an engineering manager with Mentor Graphics since 2006, responsible for the development of diagnosis and yield improvement solutions. Prior to joining Mentor Graphics, he worked at LSI Logic focusing on DPPM reduction and test based yield learning. He has co-authored more than 20 professional publications and has received the VLSI Test Symposium Best Paper Award in 2003 and the International Test Conference Honorable Mention in 2004 and 2009. In addition, he has been awarded five US patents in the area of DFT and test.

Grzegorz Mrugalski received the M.S. and Ph. D. degrees in electrical engineering from Poznań University of Technology, Poland, in 1995 and 2002, respectively. In September 2002, he joined Mentor Graphics Corporation, Wilsonville, OR. Prior to joining Mentor Graphics, he worked at the Institute of Electronics and Telecommunications of Poznań University of Technology, Poznań, Poland. Since July 2009, he is a manager of an R&D lab for Design-for-Test

products at Mentor Graphics Polska in Poznań, Poland. His current research interests include computer-aided design of digital circuits, design for testability, built-in self-test, and test compression. He co-authored over 30 technical papers in the area of test and holds 12 US and international patents. He is a member of the IEEE.

Artur Pogiel received the M.S. degree in electrical engineering from Poznań University of Technology, Poland, in 2002, and the Ph.D. degree in telecommunications from the same university in 2008. He was a teaching assistant at the Faculty of Electronics and Telecommunications of Poznań University of Technology until 2008. He is now with the Mentor Graphics Polska, Poznań, Poland, as a software development engineer. He was co-recipient of the 2009 Best Paper Award at the VLSI Design Conference. His main research interests include design for testability, built-in self-test, embedded test, and fault diagnosis. He co-authored over 10 technical papers in the area of test and holds seven US and international patents. He is a member of the IEEE.

Janusz Rajski received the M.Eng degree in electrical engineering from the Technical University of Gdańsk, Poland, in 1973, and the Ph.D. degree in electrical engineering from Poznań University of Technology, Poland, in 1982. From 1973 to 1984, he was a member of the faculty of Poznań University of Technology. In June 1984, he joined McGill University, Montreal, Canada, where he became Associate Professor in 1989. In January 1995 he accepted the position of Chief Scientist at Mentor Graphics Corporation, Wilsonville, OR. His main research interests include design automation and testing of VLSI systems, design for testability, built-in self-test, and logic synthesis. He has published more than 150 research papers in these areas and is co-inventor of 77 US and international patents. He is also the principal inventor of Embedded Deterministic Test (EDT™) technology used in the first commercial test compression product TestKompress®. He is co-author of *Arithmetic Built-In Self-Test for Embedded Systems* published by Prentice Hall in 1997. He was co-recipient of the 1993 Best Paper Award for the paper on logic synthesis published in the *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, co-recipient of the 1995 and 1998 Best Paper Awards at the IEEE VLSI Test Symposium, co-recipient of the 1999 and 2003 Honorable Mention Awards at the IEEE International Test Conference, as well as co-recipient of the 2006 IEEE Circuits and Systems Society Donald O. Pederson Outstanding Paper Award recognizing the paper on embedded deterministic test published in the *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, and the 2009 Best Paper Award at the VLSI Design Conference. Dr. Rajski was guest co-editor of the June 1990 and January 1992 special issues of *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* devoted to the 1987 and 1989 International Test Conferences, respectively. In 1999, he was a guest co-editor of the special issue of the *IEEE Communications Magazine*

devoted to testing of telecommunication hardware. He was also associate editor for *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, *IEEE Transactions on Computers*, and *IEEE Design and Test of Computers Magazine*. He has served on technical program committees of various conferences including the IEEE International Test Conference and the IEEE VLSI Test Symposium. He is a senior member of the IEEE.

Jędrzej Solecki received the B.S. degree in computer science from Poznań University of Technology, Poland, in 2010. He is currently working towards his M. S. degree at the Faculty of Computing Science of Poznań University of Technology, Poznań, Poland. His main research interests include design for testability, built-in self-test, fault diagnosis, and embedded systems.

Jerzy Tyszer received the M.Eng (Hons.) degree in electrical engineering from Poznań University of Technology, Poland, in 1981, the Ph.D. degree in electrical engineering from the same university in 1987, and Dr. Habilis degree in telecommunications from the Technical University of Gdańsk, Poland, in 1994. From 1982 to 1990, he was a member of the faculty of Poznań University of Technology, Poland. In January 1990, he joined McGill University, Montreal, Canada where was Research Associate and Adjunct Professor. In 1996, he assumed the position of Professor at the Faculty of Electronics and Telecommunications of Poznań University of Technology, Poznań, Poland. His main research interests include design automation and testing of VLSI systems, design for testability, built-in self-test, embedded test, and computer simulation of discrete event systems. He was co-recipient of the 1995 and 1998 Best Paper Awards at the IEEE VLSI Test Symposium, the 2003 Honorable Mention Award at the IEEE International Test Conference, the 2006 IEEE Circuits and Systems Society Donald O. Pederson Outstanding Paper Award recognizing the paper on embedded deterministic test published in the *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, and the 2009 Best Paper Award at the VLSI Design Conference. He has published 7 books, more than 100 research papers in the above areas and is co-inventor of 57 US and international patents. Dr. Tyszer is co-author of *Arithmetic Built-In Self-Test for Embedded Systems* published by Prentice Hall in 1997 and the author of *Object-Oriented Computer Simulation of Discrete Event Systems* published by Kluwer Academic Publishers in 1999. In 1999, he was a guest co-editor of the special issue of the *IEEE Communications Magazine* devoted to testing of telecommunication hardware. He has served on technical program committees of various conferences including the IEEE International Test Conference, the IEEE VLSI Test Symposium, and the IEEE European Test Symposium. He is a senior member of the IEEE.